

DOCUMENT RESUME

ED 138 611

TM 006 229

AUTHOR Paine, Whiton Stewart
TITLE The Utility of Goal Attainment Scaling for Educational Evaluation. Research & Evaluation Report No. 5.
INSTITUTION Crozer-Chester Medical Center, Chester, Pa. Research and Evaluation Service.
PUB DATE [Apr 77]
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977) ; Figure 8 may be marginally legible
EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.
DESCRIPTORS Educational Objectives; *Evaluation Methods; Measurement Techniques; *Program Evaluation
IDENTIFIERS *Goal Attainment Scaling

ABSTRACT

Evaluators are becoming increasingly specialized and new techniques developed in one subfield are often not well known in other subfields. Goal Attainment Scaling is a procedure originally developed to evaluate individualized treatment outcomes in mental health agencies. However, as a general evaluation procedure it also provides a flexible tool for the evaluation of educational programs. This paper describes the method, discusses past educational applications, and suggests some general utilities and limitations of Goal Attainment Scaling. (Author)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *



R & E REPORT #5
GOAL ATTAINMENT SCALING

RESEARCH & EVALUATION REPORT

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Prepared by the

Research and Evaluation Service
Community Mental Health Center
Crozer-Chester Medical Center

For copies contact:

Whiton Stewart Paine, Ph. D.
Dir., Research and Evaluation Service
Crozer-Chester Medical Center
Community Mental Health Center
1439 Upland Avenue
Chester, Pennsylvania 19013
Phone: (215) 485-7500, Ext. 641

R & E REPORT #5

The Utility of Goal Attainment Scaling
for Educational Evaluation

Whiton Stewart Paine
Community Mental Health Center
Crozer-Chester Medical Center

AERA Session 28.10
4/8/77 New York

Running Head: Goal Attainment Scaling

Abstract

Evaluators are becoming increasingly specialized and, partially as a result, new techniques developed in one subfield are often not well known in other subfields. Goal Attainment Scaling is a procedure originally developed to evaluate individualized treatment outcomes in mental health agencies. However, as a general evaluation procedure it also provides a flexible tool for the evaluation of educational programs. This paper describes the method, discusses past educational applications, and suggests some general utilities and limitations of Goal Attainment Scaling.

The Utility of Goal Attainment Scaling
for Educational Evaluation

Evaluation is a rapidly growing interdisciplinary field which is diversifying into a cluster of subfields. As professionals employ a common body of concepts in different types of applied settings, new techniques are developed. Unfortunately, the dissemination of these advances is often slow. This paper is an attempt to present a technique which was developed by mental health evaluators, but which is relevant to the needs of evaluators working with educational problems.

Goal Attainment Scaling (GAS) is a relatively simple procedure for operationalizing goals and measuring their attainment. It involves setting goals and defining five levels of expected attainment or non-attainment. The degree of attainment is measured on a follow-up basis. Most of the work on GAS has involved the measurement of individual treatment outcomes in mental health agencies. Recently, there have been a number of interesting educational applications of GAS, and the basic concepts have been extended in order to create a general evaluation tool. Finally, some of the difficulties and limitations of the procedure are becoming known. Each of these points is discussed below, beginning with a brief description of the Goal Attainment Scaling procedure.

Basic Procedure

The basic procedure for Goal Attainment Scaling was first published in a 1968 article by Thomas J. Kiresuk and Robert E. Sherman. Respectively, they were a clinical psychologist and a bio-statistician who were interested in evaluating the clinical programs of a mental health agency. By 1976, their work had led to the development of the Program Evaluation Resource Center, which publishes Evaluation magazine; and to an increasing network of agencies utilizing the technique. A recently published bibliography (Garwick, McCree & Brintnall, 1976) lists 170 citations on GAS, most of which are dated since 1970.

Goal Attainment Scaling is not a specific procedure, rather it consists of variations on the basic set of steps for constructing and using a Goal Attainment Follow-Up Guide. These steps are summarized in a number of recent publications by Geoffrey Garwick (1976, 1975 a & b), and his analysis is the basis for this section. A sample guide is reproduced on pages 5a and b alongwith the scoring formula.

First, areas of concern or problem areas are selected and a set time for the follow-up assessment is established. Optionally, these areas may be weighted according to their importance.

Specific goals are then selected for each problem area. These are observable indicators, and they correspond to the way many educational evaluators use the term "objective". Again, a specific time period for follow-up is defined.

There are five levels of possible goal attainment ranging from the best to the worst outcome thought likely. These are defined starting with a realistic, pragmatic objective, the "expected outcome" at the end of the follow-up period. Then the "much less" and "much more" than

expected levels are defined followed by intermediate "less" and "more" levels. All four levels do not need to be defined. The intent of this procedure is to create a five point scale which is a "judgemental transformation of the treatment outcome, into, approximately, a random variable with a mean zero and variance one (sic)" (Kiresuk & Sherman, 1968, p. 447). Each goal is weighted according to the weights set for the entire problem area, or separate goal weights can be used. If a change measure is of interest, the present level of functioning of the client may be noted on a separate sheet.

The Goal Attainment Follow-Up Guide should be carefully reviewed. Then, after the follow-up interval, the guide is used to assess the final level of goal attainment. The set of final level scores can be transformed into a single composite scale with a mean of 50 and a standard deviation of 10 by using a formula developed by Kiresuk & Sherman (1968). Published tables are available to facilitate this transformation. (Garwick & Brintnall, 1973).

Garwick has summarized much of the reliability and validity work on GAS and has prepared a set of guidelines for GAS (1975a). In this collection, he promulgated the following ten commandments for construction of the follow-up guide. They are (p. 27):

1. Include at least three scales on a follow-up guide except in dire straits, and even then have at least two scales.
2. Have at least three levels filled-in for each scale. One filled-in level should be the expected level, and there should be at least one filled-in level on each side of the expected level.
3. Include only one problem or variable on each scale.
4. If any of the scales on a follow-up guide are weighted, all scales should be weighted.

5. The client's behavior at intake may be equivalent to any of the five levels. Avoid terms like "better than when treatment began."
6. Avoid variables which are too general or vague to be accurately scored at a follow-up interview.
7. The levels on a scale should not overlap each other.
8. If the information needed for a scale's follow-up scoring is to be obtained from a source other than the client, the special source of information should be listed.
9. There should be no "expected" levels which are so high that there is no possible "better than expected" level. There should be no "expected" goals which are so low that there is no possible "less than expected" levels. There should usually be a possible outcome for all levels of the scale even if some are left blank.
10. Avoid having two blank levels adjacent to each other on a scale.

Almost every element of this basic GAS format can be modified to meet specific program or treatment needs. In their original article, Kiresuk and Sherman (1968) presented a highly structured evaluation system that operated largely independently of the clinical process. Staff other than the treatment personnel were responsible for preparing the sheets and performing the follow-up. Most users of GAS have varied widely from the Kiresuk and Sherman (1968) approach and have sacrificed evaluation rigor for other goals. The effects of these changes on the measurement of programs is not known but they have significantly increased the flexibility of the approach. This flexibility is demonstrated in the range of published educational applications of GAS.

Figure 8: Scales Completed, with Original Levels and Follow-up Levels Indicated

NOTE: * is original Level
x is follow-up Level
GAS score is 54.2
GAS gain score is +7 levels

GOAL ATTAINMENT FOLLOW-UP GUIDE						
LEVELS OF PREDICTED ATTAINMENTS	SCALE HEADINGS AND SCALE WEIGHTS					
	SCALE 1: FIRESETTING (weight ₁ = 37)	SCALE 2: GRADES (weight ₂ = 37)	SCALE 3: SELF-CONCEPT (weight ₃ = 50)	SCALE 4: YELLING IN CLASS (weight ₄ = 10)	SCALE 5: Singing (weight ₅ = 8)	SCALE 6: SEXUALITY (weight ₆ = 20)
most unfavorable outcome thought likely (-2)	2 or more fires set.*	0 average or less in latest quarter.	Can list only 1 good accomplishment when asked.	Teacher reports daily yelling during study time in last month.	Refuses to sing in music, says he's tone deaf.	Caught writing obscenities on home or school walls 2 or more times during year. *.
less than expected success (-1)	1 fire set.	0+ or C-.		Yelling 4 days per week.	Refuses to sing, but says its for reasons other than tone deaf- ness.	10 to 19 times. *
expected level of success (0)	He reports no fires set in next year but 10 urges per day to set fires.	C.	List 4 to 6.	Yelling 3 days per week.	Sings with class at least once during last quarter.	5 to 9 times.
more than expected success (+1)	No fires set, but 5 to 10 urges per day.	C+.		Yelling 1 or 2 days per week.	Sings with class 2 to 5 times.	2 to 4 times.
most favorable outcome thought likely (+2)	No fires set, but four or less urges per day.	B- or better.	Lists 9 or more.	Yelling less than 1 day per week.	Sings with class 6 or more times.	1 or less.

The formula for GAS score is:

$$10 \leq \sum_{i=1}^n W_i X_i$$

$$G.A.S. = 50 + \sqrt{(1-\rho) \sum_{i=1}^n W_i^2 + \rho \left(\sum_{i=1}^n W_i \right)^2}$$

X_i = Outcome score on i^{th} scale of follow-up guide

W_i = Relative weight attached to the i^{th} scale

ρ =(rho) = Weighted intercorrelation of scales (usually set at .3)

n = Number of scales on the follow-up guide

Educational Applications

Goal attainment scaling has been primarily employed to evaluate individualized treatment outcomes. More recently there has been an increased interest in GAS as a general tool for program evaluation and this is discussed in the next section. The limited available literature on educational applications of GAS reflects this changing emphasis; with studies emphasizing individual outcomes, adapting standardized outcomes to individual differences, and attempting to evaluate entire programs.

These studies are presented in some detail because they illustrate quite different attempts to employ GAS in situations familiar to educational evaluators.

In general, little work has been done on utilizing GAS with children (Gillis & Paine, 1977). However, two reports are available on the technique as it has been applied in programs for children with special educational needs. A Minnesota project on the early identification of children with learning disabilities utilized a fairly standard evaluation design in which children in experimental and control schools are compared on teacher ratings and on standardized tests (Hegion, Fish & Grace, 1974). In addition to these commonly employed measures, the program felt that they needed a way of "measuring and monitoring the progress of each child on an individual basis, and in a standardized manner" (Hegion, Fish & Grace, 1976, p. 20).

Project staff worked with teachers in defining each student's goals for the follow-up guides. Most goals were in specific areas of basic academic skills, but social and classroom behavior goals were established for some children. The project also defined some administrative goals to assess the delivery of services to the teachers.

Commonly, GAS scores have a mean of 50 and a standard deviation of 10. The data on the two schools indicated a higher mean and a smaller standard deviation. This suggests that the teachers had underestimated the performance of children in the project. Rank correlations across the two schools were high and the means differed slightly between schools.

Unfortunately, this report does not contain details on the behavioral and administrative GAS procedures which were also used in this study. However, it is useful because it lists, in an appendix, the basic academic skills and goals used in the study. These eight pages of goals give the reader a good idea of how GAS can be used to monitor the academic performance of elementary school students.

A very different application was reported by Miller (1975) in his description of a Canadian remedial nursery school for multi-handicapped children. In this setting, a variety of approaches were incorporated into a records system which includes both educational and treatment goals. Their "Problem and Goal Evaluation System" is especially interesting because it adapts aspects of GAS to the demands of a heavily computerized records environment.

After admission, children are diagnosed in terms of 115 specific problems or sub-problems and this leads to a computer generated problem profile. Goal sheets are used to record the goals for each problem and the child's present level is rated and the future week for follow-up is set. Up to nine goals can be specified for each problem or sub-problem. Follow-up sheets are generated automatically and the child is assessed in terms of nine categories of goal attainment. Periodically a progress record is generated and ratings of the performance of staff are also routinely produced. This is a highly automated goal oriented system

which is designed to be simultaneously used for educational, clinical, administrative and research purposes.

An innovative use of GAS in higher education is reported by McGahie & Menges (1975). As part of an Introductory Psychology class, students were required to complete individual behavioral self modification projects. The problem of interest was to develop a technique which would allow both individual and group assessment of behavioral outcomes. This is the classic problem of individualized learning versus normative assessment.

The authors report a 72% success rate for 191 students over two years of classes. Behavioral and emotional change correlated moderately highly both at the end of the project and at an eight week follow-up. Correlations across time were considerably lower. McGahie & Menges (1975) feel that the GAS has adequate internal consistency, low stability, adequate face and content validity, and some predictive validity in their application.

This is an interesting study because it both merges GAS and concepts drawn from behavior modification, and it also provides techniques for assessing self-directed study and similar educational formats.

Goal Attainment Scaling has proven useful in the evaluation of adult and continuing education formats. Canfield and Klewer (1975) have prepared a detailed conference evaluation manual which has been published as part of the five part NIMH sponsored set of Resource Materials for Community Mental Health Program Evaluation (Hargreaves, Attkisson, Siegel, McIntyre & Sorensen, 1974, 1975). This manual was used to evaluate a number of conferences but the outcomes are not presented in this report. This approach combines Likert and GAS items, and both conference and

participant goals can be accommodated. Three levels of attainment are used.

The Program Evaluation Resource Center used GAS to evaluate some of their workshops (Garwick & Vanderpool, 1975). The Workshop Committee prepared ten scales in advance, and one was discarded because of "incoherency". Data on six scales is presented indicating higher than expected attendance and perceived utility, expected levels of content mastery, and lower than expected outcomes on the number of written comments received and number of conference participants returning the follow-up questionnaire. Goals were not individualized.

A very detailed analysis of the impacts of a communications skills training format for family practice residents is reported by Carlson (1974). The exhaustiveness of the analysis is probably not warranted by the small number of residents involved, but this report and the previous one are useful in suggesting ways in which standard questionnaire and GAS results can be combined in an evaluation report.

Benedict's (1974) analysis of a twelve week small group methods seminar also suffers from a small sample size. This paper, however, is something of a model attempt to use GAS to integrate different sources of data and widely differing goals into an exhaustive, recommendation oriented evaluation of a training event. The training format was based on a specific model and organized around goals negotiated with participants in advance. A range of attitudinal and behavioral measures were employed to evaluate both process and outcome goals. Benedict (1974) provides considerable detail on the rationale for each goal and on how recommendations were generated in response to the differing levels of goal attainment. Some long term impacts of the training on the agency involved are also discussed.

These educational applications of GAS are fairly representative of the general clinical literature on the technique. Each is interesting because it suggests a different set of possible future evaluation studies in elementary schools, higher education and continuing education. A separate set of applications is suggested by the growing literature on GAS as a general tool for evaluating all levels of a program or project. This type of evaluation is briefly discussed in the Hegion, Fish & Grace (1974) and Miller (1975) studies but other workers have provided more detail on this use of GAS. Some of these reports are discussed in the next section as suggestions for possible future applications in educational evaluation.

Use in Program Evaluation

While GAS has been utilized most often as an individualized assessment procedure, it can also be effectively employed to assess entire programs. Aside from the obvious application of using goal attainment scores as data points in an evaluation of treatment outcomes, there are numerous other ways in which GAS can be useful. Kiresuk & Sherman (1968, p. 447) stressed "the capacity of the scaling technique to utilize any form of objectively determinable event," and this characteristic opens up many opportunities for evaluators of education programs.

For example, if the comparative assessment of individual staff is of interest, a study in Missouri provides a possible model (Barron & Bode, 1974). The task was to assess the performance of regional program development specialists who were employed to facilitate the development of a coordinated network of alcohol services throughout the state. Barron & Bode (1974) used a technique similar to the Delphi format to obtain consensually valid goals from the supervisors, recipients and workers in this governmental program.

These same individuals used the final set of 16 scales to rate the performance of the regional specialists. The averages for different sets of standardized goal attainment scores varied widely and only those relating to administration reflected an expected level of attainment. Goal sets reflecting the development of, and the coordination of programs were attained at a less than expected level. As is often the case in evaluation, goals with higher weights, reflecting their greater importance, were not attained as well as were less important goals. An analysis of factors effecting goal attainment found that amount of travel and amount of coordination effort required were both negatively coordinated

with attainment. The effects of such factors should be analyzed in all application of GAS for program evaluation.

A second example of the use of GAS as a monitoring tool by governmental agencies is reported by VanHoudnos and Williams (1974). In 1973, the state of Illinois began a management by objectives program for all state agencies. One regional office of the Department of Mental Health utilized GAS to satisfy the mandate for MBO for both the region and for two facilities monitored by the region. The regional office and one facility were able to implement GAS, the second facility was in the midst of a 30% staff reduction and had other priorities. Unfortunately this report is too short to fully assess the project, but it suggests that GAS can be successfully integrated into MBO formats.

A county social service agency (Ames, 1976) identified five basic areas of staff functioning and set goals for staff in each area. The staff as a whole defined the various levels of expectation and their performance is assessed periodically. In this setting the results are used for salary and promotion decisions. Ames' (1976) preliminary report indicates high levels of staff satisfaction with this approach.

One crisis intervention center used an internal planning committee to establish eighteen program goals and their levels (Lund, 1976). The selected goals were quite heterogeneous and involved different types of information. Many were scaled independently of the administrative staff. On follow-up, the overall GAS score was 64.08 which is over one standard deviation higher than the "expected" level score of 50, with the more objective scales tending toward higher scores. This initial project has led to an on-going evaluation system for the Center.

The use of GAS in a behavior modification oriented evaluation system was reported in a recent edition of Evaluation magazine (King, Austin, Liberman & DeRisi, 1974). The Behavior Analysis and Modification Project is investigating the use of GAS in evaluating behavior therapy in a CMHC. The staff also set up nine goals related to their research, training, evaluation and dissemination activities. Staff reported many of the usual advantages of using any goal oriented technique, including planning, monitoring and feedback related benefits.

Paine (1977) used GAS in a situation where there were no program goals. Telephone follow-up satisfaction studies were being done on clients who had left different treatment services of a comprehensive CMHC. Appropriate external standards are not available for this type of exploratory work, and the CMHC will not set internal standards until the studies are done. Therefore, the procedure was adopted of taking five key Likert or yes/no items from the interview form, and asking service directors to indicate the expected levels of client response.

The unit directors were directed to adjust these expectations to reflect client characteristics, treatment philosophy of the service, staff characteristics, and available resources. The GAS sheets are then used to compute aggregate attainment scores and to structure written and verbal feedback to the units.

A side benefit of this procedure is that client satisfaction surveys generally have a positive bias. Thus, the service directors expectations are likely to be low and the resultant feedback will probably be positive. This will aid later evaluation work on these units as well as facilitate a realistic readjustment of expectations before CMHC wide standards are promulgated.

The most extensive and best integrated attempt to use GAS as a comprehensive evaluation procedure appears to be the system under development at the Hennepin County Mental Health Service (Baxter & Beaulieu, 1974; Kiresuk, 1973; Kiresuk & Lund, 1975). Baxter & Beaulieu (1974) summarize the overall model and its relationship to the concepts used in goal attainment scaling and management by objectives. Procedures have been developed for the assessment of client outcomes, process evaluation and for the monitoring and evaluation of the entire program.

This work is of particular interest because it is beginning to generate data on large groups of clients and also is providing opportunities for different types of studies testing the characteristics and the organizational impacts of GAS. By stressing the measurement of both efficiency and effectiveness, Baxter & Beaulieu (1974) are moving toward a comprehensive integrated internal evaluation system.

Again, this is a set of reports which is suggestive of a wide range of possible GAS applications for program evaluation. Little is known, however, about the effects of non-clinical applications on the utility, reliability and validity of GAS. There also are additional questions which can be raised about the entire approach.

Limitations

Goal attainment scaling is certainly not a panacea and there can be a variety of problems with the techniques. Despite its apparent simplicity, GAS is not an easy system to implement nor is it an inexpensive system to operate. Individuals who are not knowledgeable about behavioral goal setting may have difficulty understanding the importance of careful specification of goal levels (Paine & Gillis, 1976). Follow-up contacts are often expensive, and with some populations, considerable loss of data is to be expected (Paine, 1977). Paine (1976) analyzed three attempts to implement GAS using the A VICTORY model of planned change (National Institute of Mental Health, 1973; Davis & Salison, 1975). Difficulties were identified in each of the eight conceptual areas of the model, which suggests the range of implementation problems that can be encountered.

The National Institute of Mental Health investigated the utility of GAS as a possible technique for monitoring evaluation research grants (Windle, Schwitzgebel, Zusman & McPhee, 1974). This effort was unsuccessful because redesign activities and implementation problems in these time limited projects often invalidated the original set of goals. Goal attainment scaling is most likely to be successful in fairly stable organizational environments and may not be useful in many short term projects or in programs undergoing rapid change.

There is little information on the relative costs and benefits of modifying Kiresuk's (1968) original, rather rigorous, format. Such information is needed if comparisons across programs are to be made. Ideally, GAS would be a very useful tool for directly comparing diverse types of programs since a standardized goal attainment score could be computed for any set of program goals. A program would be deemed successful if it ob-

tained a GAS score of 50 plus or minus some preset percentage of the standard deviation of 10. Unfortunately, computing the GAS score immediately raises some serious problems of interpretation.

Depending upon how the expected level is defined, an expected GAS score of 50 may be difficult to interpret across programs. This is a problem reflecting the different ways in which the expected level is defined. A program may define itself as successful if it achieves either a specific number (e.g., admissions) or if performance falls somewhere within a range of values or both depending on the goal. To further complicate this problem; process, short-term and long-term goals may be mixed together in the same outcome analysis. Different types of goals should be grouped together and analyzed separately if the final GAS score is to be meaningful.

Secondly, a score above or below 50 can almost always be interpreted in alternative ways. Either the program was more, or less, successful than expected or the initial goal levels reflected inappropriate expectations. This is a common problem with goal oriented approaches and suggests that some form of external review is probably necessary to validate the appropriateness of the goal levels.

The latter problem is not trivial. The GAS formula reproduced on page 5b is designed to yield a normal distribution of scores with a Mean of 50, a standard deviation of 10 and a range of a little over 60 points. While the distribution of the Hennepin County data does have the expected mean of 51.9 and a standard deviation of 10.1 (Kiresuk & Lund, 1975) other available studies have deviated markedly from these values. In fact, the reported GAS scores of under 40 or over 60 may be more common than those that fall within one standard deviation of the mean. One reviewer

(Ellsworth, 1975) considers this problem serious enough to disallow GAS as a general evaluation technique.

Other psychometric properties of the approach are also not well understood. Garwick (1974b) has pointed out that the traditional conceptions of reliability and validity must be considerably modified to test this type of scaling procedure.

Kiresuk and his colleagues are beginning to investigate these complex reliability and validity questions (Garwick, 1974 a, b, c; Mauger, Audette, Simonini & Stolberg, 1974; Sherman, Baxter & Audette, 1974). To date, the studies that have been done indicate that measures of internal consistency and inter-rater reliability and of the relationship of individual scales to the total score are generally around .6 - .7, and that the technique is less stable over time than such procedures as the MMPI and the therapist's global ratings.

A content analysis study (Grygelko, Garwick & Lampman, 1976) indicates that the independent guides of therapists and interviewers were similar in their relative emphasis on the twenty most common generic goal areas but that they frequently differ markedly on the use of the different goals with individual clients. The latter finding is important because the approach to content validity maintains that, in the absence of appropriate alternative measures, content validity equals reliability, since a pair of raters is assumed to be randomly chosen from the universe of all pairs of raters setting goals with the client.

Garwick (1974 a, c) is also building a complex "nomological network" to begin the analysis of the construct validity of GAS. The construct of interest is treatment outcome and this analysis yields a number of interesting predictions. Since the GAS procedure is individualized it should

correlate weakly but positively with non-individualized standardized outcome measures. Correlations with other individualized rating procedures such as client satisfaction should be somewhat higher but still moderate. Finally, GAS scores should not reflect either client or treatment differences since such differences would be taken into account in the construction of the follow-up guides. Garwick's surveys (1974 a, c) indicate that most, but not all, of the available data supports these predictions. Much more work needs to be done to resolve some of these issues related to GAS as a measurement tool.

Summary

This brief report suggests the possible range of applications of Goal Attainment Scaling. Educational evaluators have an obvious need for a technique which allows for the assessment of both individualized and general program goals. Since so much of evaluation is goal oriented, GAS should be considered in any such study as a possible technique. Similarly, it appears compatible with studies employing management by objectives and, possibly, cost-benefit analyses. If the reader is interested in any of the applications described, or in the general method, most of the reports cited can be obtained from the Program Evaluation Resource Center or appear in available journals. The address of PERC is 501 Park Avenue South, Minneapolis, Minnesota 55415. Joan Brintnall is in charge of publications and the phone number is (612) 348-7811.

The purpose of this paper is to facilitate the access of educational evaluators to this general evaluation technique. One of the major problems with GAS is that so much of the past work has been associated with PERC and this. Studies of GAS done by outside evaluators are needed to test the real benefits and limitations of the approach. The author would appreciate copies of any such studies.

REFERENCES

- Ames, R. Staff performance evaluation with Goal Attainment Scaling in a county social service agency. in J. Brintnall & G. Garwick (Eds.), Applications of Goal Attainment Scaling, Minneapolis: Program Evaluation Resource Center, 1976.
- Barron, N. & Bode, E. Goal Attainment Scaling and regionalized alcohol program development. In G. Garwick (Ed.), Goal Attainment Review Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.
- Baxter, J. W. & Beaulieu, D. E. Evaluation of the adult outpatient program, Hennepin County Mental Health Service. P.E.P. Report 1969-1973 (Chapter 9). Minneapolis: Program Evaluation Resource Center, 1974.
- Benedict, W. Utilizing Goal Attainment Scaling to evaluate a staff development program. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.
- Canfield, M. & Kliwer, D. Conference evaluation manual. In Hargreaves, W. A., Attkisson, C. C., Siegel, L. M., McIntyre, M. H. & Sorensen, J. E. Resource materials for community mental health program evaluation. Part IV - Evaluating the effectiveness of services. Rockville, Md., National Institute of Mental Health, 1975.
- Carlson, G. Communication skills training in a family practice residency program. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.
- Carter, D. E. & Newman, F. L. A client-oriented system of mental health service delivery and program management - A workbook and guide. (DHEW Publication No. (ADM) 76-307, stock number 017-024-00523-1) Washington, D. C.: U.S. Government Printing Office, 1976.

- Cline, D. W., Rouzer, D. L. & Bransford, D. Goal-Attainment Scaling as a method for evaluating mental health programs. American Journal of Psychiatry, 1973, 130, 105-108.
- Davis, H. R. & Salasin, S. E. The utilization of evaluation. In E. L. Struening & M. Guttentag (Eds.), Handbook of evaluation research (Volume 1). Beverly Hills, Sage Publications, 1975.
- Ellsworth, R. B. Consumer feedback in measuring the effectiveness of mental health programs. In E. L. Struening & M. Guttentag (Eds.) Handbook of evaluation research (Volume 2). Beverly Hills, Sage Publications, 1975.
- Garwick, G. A construct validity overview of Goal Attainment Scaling. P.E.P. Report 1969-1973 (Chapter 5). Minneapolis: Program Evaluation Resource Center, 1974. (a)
- Garwick, G. An introduction to reliability and the Goal Attainment Scaling methodology. P.E.P. Report 1969-1973 (Chapter 3). Minneapolis: Program Evaluation Resource Center, 1974. (b)
- Garwick, G. Some recent findings on the use of goal-setting in human service agencies. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974. (c)
- Garwick, G. Guidelines for Goal Attainment Scaling. Minneapolis: Program Evaluation Resource Center, 1975. (a)
- Garwick, G. A quick review of the Goal Attainment Scaling process. In G. Garwick & J. Brintnall. Proceedings of the Second Goal Attainment Scaling Conference. Minneapolis: Program Evaluation Resource Center, 1975. (b)

- Garwick, G. The rudiments of Goal Attainment Scaling. In J. Brintnall & G. Garwick (Eds.), Applications of Goal Attainment Scaling. Minneapolis: Program Evaluation Resource Center, 1976.
- Garwick, G., McCree, S. & Brintnall, J. Bibliography on Goal Attainment Scaling and associated methodologies. Minneapolis: Program Evaluation Resource Center, 1976.
- Grygelko, M., Garwick, G. & Lampman, S. Findings of content analysis: 1) Patterns of use 2) Reliability. In J. Brintnall & G. Garwick (Eds.), Applications of Goal Attainment Scaling, Minneapolis: Program Evaluation Resource Center, 1976.
- Hegion, A., Fish, E. & Grace, S. Goal Attainment Scaling: An alternative for measuring academic progress. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.
- King, L. W., Austin, N. K., Liberman, R. P. & DeRisi, W. J. Problems in review-Accountability, like Charity, begins at home. Evaluation, 1974, 2, 75-77.
- Kiresuk, T. J., Goal Attainment Scaling at a county mental health service. Evaluation, Special Monograph No. 1, 1973, 12-18. Reprinted in W. A. Hargreaves, C. C. Attkisson, L. M. Siegel, M. H. McIntyre & J. E. Sorensen. Resource materials for community mental health program evaluation, Part IV - Evaluating the effectiveness of services. Washington, D. C.: National Institute of Mental Health, 1975.
- Kiresuk, T. J. & Lund, S. H. Process and outcome scaling using Goal Attainment Scaling. In J. Zusman & C. R. Wurster (Eds.), Program evaluation: Alcohol, drug abuse and mental health services. Lexington, Massachusetts: Lexington Books, 1975.

- Kiresuk, T. J. & Sherman, R. E. Goal Attainment Scaling: A general method for evaluating comprehensive community mental health programs. Community Mental Health Journal. 1968, 4, 443-453.
- Lund, S. Crisis intervention center: Administrative evaluation. In J. Brintnall & G. Garwick (Eds.), Applications of Goal Attainment Scaling. Minneapolis: Program Evaluation Resource Center, 1976.
- Mauger, P., Audette, D., Simonini, C. & Stolberg, A. A study of the construct validity of Goal Attainment Scaling. In G. Garwick (Ed.) Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.
- McGaghie, W. C. & Menges, R. J. Assessing self directed learning. Teaching of Psychology. 1975, 2, 56-59.
- Miller, G. H. A problem and goal oriented evaluation system. In B. Willer, G. H. Miller & L. Cantrel. Information and feedback for evaluation. Toronto: The Communications Department, York University, 1975.
- National Institute of Mental Health. A distillation of principles on research utilization (Volume 1). DHEW Publication No. (HSM) 73-9148 GPO. Washington, D. C.: U.S. Government Printing Office, 1973.
- Paine, W. S. Some problems in implementing Goal Attainment Scaling. An A VICTORY analysis. Paper presented at the Conference on Clinical Records Systems, Chester, Pa., November 19, 1976.
- Paine, W. S. Routinizing periodic client satisfaction surveys in a comprehensive community mental health center. Paper presented at the National Council of Community Mental Health Centers Annual Meeting, Atlanta, March 1, 1977.

Paine, W. S. & Gillis C. Implementing innovative clinical records in a comprehensive community mental health center. In The National Project on Education for Management in Social Welfare. Management in social welfare: Casebook studies. Philadelphia: School of Social Work/Wharton School, 1976.

Sherman, R. E., Baxter, J. W., & Audette, D. M. An examination of the reliability of the Kiresuk-Sherman goal attainment score by means of components of variance. P.E.P. Report 1969-1973 (Chapter 4).

Minneapolis: Program Evaluation Resource Center, 1974.

Swanstrom, W. & Hedberg, L. Broadening the use of Goal Attainment Scaling in residential treatment. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.

VanHoudnos, H. & Williams, L. Program evaluation, management by objectives, and the Goal Attainment Scaling format. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.

Windle, C., Schwitzgebel, R. L., Zusman, J. & McPhee, C. Goal Attainment Scaling for research monitoring. In G. Garwick (Ed.), Goal Attainment Review (Volume 1). Minneapolis: Program Evaluation Resource Center, 1974.